TEN

Perception

It is not surprising, noting the visual bias of modern Western culture, that the psychology of aural perception has been comparatively neglected. Much of the work done has been concerned with binaural hearing and sound localization—which also has largely to do with space. Quite a lot of work has been done on masking (covering one sound by another) and some has been done on auditory fatigue (the effect of prolonged exposure to the same sound); but taken as a whole such researches leave us a long way from our goal, which would be *to determine in what significant ways individuals and societies of various historical eras listen differently.*

Thus it is inconceivable that a music or soundscape historian should get quite the same thrill out of the preparatory work the laboratories have provided as that which has stimulated art historians such as Rudolph Arnheim and E. H. Gombrich, whose work owes such a heavy debt to research in the psychology of visual perception. In the work of men like these it has begun to be possible to comprehend the history of vision, at least in the Western world. The soundscape historian can only speculate tentatively on the nature and causes of perceptual changes in listening habits and hope that psychologist friends may respond to the need for more experimental study.

Figure and Ground It is indeed possible that some terms employed in visual perception may have equivalents in aural perception. At least they are probably worth careful examination. For instance, a phenomenon like irradiation—by which a brightly illuminated area seems to spread—does seem to have an analogy in that a loud sound will appear to

be longer than a quiet one of equal duration. It is still not clear whether a term like *closure*—which refers to the perceptual tendency to complete an incomplete pattern by filling in gaps—can be applied to sound with anything like the confidence it has stimulated in visual pattern perception, though experiments in phonology show that for language at least there are striking parallels.

Throughout this book I have been using another notion borrowed from visual perception: figure versus ground. According to the gestalt psychologists, who introduced the distinction, figure is the focus of interest and ground is the setting or context. To this was later added a third term, *field,* meaning the place where the observation takes place. It was the phenomenological psychologists who pointed out that what is perceived as figure or ground is mostly determined by the field and the subject's relationship to the field.

The general relationship between these three terms and a set I have been employing in this book is now obvious: the figure corresponds to the signal or the soundmark, the ground to the ambient sounds around it—which may often be keynote sounds—and the field to the place where all the sounds occur, the soundscape.

In the visual figure-ground perception test, the figure and ground may be reversed but they cannot both be perceived simultaneously. For instance, looking into the clear water of a pond, one may perceive one's own reflection or the bottom of the pond, but not both at the same time. If we are to pursue the figure-ground issue in terms of aural perception, we will want to fix the points when an acoustic figure is dropped to become an unperceived ground or when a ground suddenly flips up as a figure—a sound event, a soundmark, a memorable or vital acoustic experience. History is full of such examples and this book is revealing a few of them.

Whether a sound is figure or ground has partly to do with acculturation (trained habits), partly with the individual's state of mind (mood, interest) and partly with the individual's relation to the field (native, outsider). It has nothing to do with the physical dimensions of the sound, for I have shown how even very loud sounds, such as those of the Industrial Revolution, remained quite inconspicuous until their social importance began to be questioned. On the other hand, even tiny sounds will be noticed as figures when they are novelties or are perceived by outsiders. Thus Lara notices the noise of the electric lights in Moscow as soon as Pasternak moves her in from the country (*Doctor Zhivago*) or I notice the scraping of the heavy metal chairs on the tile floors of the Paris cafés each time I visit that city as a tourist.

The terms *figure, ground* and *field* provide a framework for organizing experience. As useful as they may be, it would be injudicious to presume that they alone could lead to the goal announced at the beginning of the chapter, for they are themselves the product of one set of cultural and perceptual habits, one in which experience tends to be organized along

perspective lines with foreground, background and distant horizon. How accurately they may apply to another society, remote from this one, is the big question we want answered.

Sonological Competence The psychologist studies the processes of perception; he does not attempt to improve them. But to run his tests he must assume some competence on the part of his subjects. As a teacher of music, my instinct tells me why so little has been accomplished to date. To report one's impressions of sound one must employ sound; any other method will be spurious. Just as we accused acousticians of playing sound false by turning it into pictures, so we accuse psychologists of playing it false by turning it into stories. This is the limitation of sound-association tests where listeners are asked to describe their impressions of taped sounds in free-association narratives. Whatever the purpose of such tests, it can hardly be to provide a description of perception. The only way to check perceptions is to devise routines by which listeners can reproduce exactly what they hear. This is why the ear training exercises of music are so useful. The dancing of the tongue in onomatopoeic mimesis is another way to check perceptions. As part of the ear cleaning program, I devised many exercises of this sort; for instance, imitate with your voice the sound of a shovel digging into sand, then into gravel, then clay, then snow. This exercise is partly memory work, partly vocal facility. Matching another person's voice, say in the repetition of a name, is another exercise designed to improve the competence of subjects for acoustic reportage.

In Chapter Two I noted how different languages have special onomatopoeic expressions for familiar animals, birds or insects. Aside from the phonetic limitations of language, the obvious differences in such words *must indicate something* about the manner in which the same sounds are heard variously by separate cultures—or is it that the animals and insects speak dialects?

Impression is only half of perception. The other half is expression. Uniting these is intelligence—accurate knowledge of perceptual observations. With impression we accommodate the information we receive from the environment.* Impression draws in and orders; expression moves out and designs. Together these activities, and perhaps some others about which we are as yet less certain, make up what Dr. Otto Laske has called "sonological competence." Laske points out that sonological competence does not result from the mere reception of sensory information. "If that were so, (psycho)acoustic knowledge would be sufficient for design, but it isn't. The difference between psychoacoustic knowledge and sonological competence is exactly the difference between a 'knowledge of, or about' and a 'knowledge-to-do,' i.e., between a knowledge of sound properties and a capability for designing." Laske insists that sonological competence

*Piaget calls these two complementary aspects of perception "accommodation" and "assimilation," but I prefer the outgoing suggestion of "expression."

applies to the most rudimentary level of perception, and as such it lies at the base of all deliberate attempts at soundscape design.

It is certainly possible that some societies possess better sonological competence than others. The evidence of this book makes it more than an assumption that this was the case when the ear was more important as an information gatherer, and the elaborate earwitness descriptions in works like the Bible and *The Thousand and One Nights* suggest that they were produced by societies in which sonological competence was highly developed. By comparison, the sonological competence of Western peoples today is weak. We have ignored our ears, hence the noise pollution problem. But in addition to our ears and voices we have today an instrument which can be used to assist in reclaiming the abilities of aural discrimination—I mean the tape recorder. With this device sounds can at last be suspended, dissected, intimately investigated. More than that, they can be synthesized and it is in this that the full potentiality of the tape recorder is revealed as an instrument uniting impression, imagination and expression. The tape recorder can synthesize sounds impossible for the voice. Take, for instance, an earthquake. The best description of one I have ever encountered is that of a radio sound effects technician.

> Time after time I have heard this item portrayed by a sudden welter of earth-shattering sound and ear-splitting screams. This is way off the mark. The earthquake effect is done in four separate parts, with a few seconds pause between each. Start with a low, shuddering rumble, bring up the gain slowly, hold for a second or two, then drop it back almost to zero. Make the sound itself by shaking two rubber balls around in a cardboard box and recording the sound at double-speed or, if you are able to do so, recording at 15 ips and playing back at 3¾ ips. Having recorded the first part of the "quake" (or "prelude" as it is known), follow on with one or two isolated crockery-smashes and mix-in once more to the rumbling effect, louder this time.
>
> Now bring in a sudden sliding, crashing sound, with a tearing metallic "ring" about it. This can be achieved by dropping a quantity of small stones on to the sloping lid of a cardboard box. The lid should be held about a foot above the table surface with a glass jam-jar (lying on its side) at the lower end of the slope. The sound sequence, thus, is that the stones strike the lid of the box, slide down its surface and strike against the side of the jam-jar before coming to rest on the table top. Record the sound at absolute maximum gain. Double-speeding may improve the item still further by both lengthening the sound and giving it a "heavier" quality. Lastly, fade in the rumbling noises once more, hold, then fade to zero.
>
> Incidentally, a most uncanny yet effective impression of brooding silence can be obtained between the individual portions of activity by recording *very* faintly, the sound of distant voices alone. "Panic" noises such as screaming and shouting, if desired, are best recorded

behind the third "falling-debris" section which may be superimposed over it.

I have often discovered in teaching that one of the best ways to press students into an examination of their perceptions is to set them similar exercises in sound synthesis using tape recorders. It is then that ignored or carelessly perceived features of a total sound complex become immediately conspicuous.

Music as the Key to Aural Perception Any investigator of the world soundscape would benefit from a knowledge of the history of music. It provides us with a large repertoire of sounds—in fact, the largest repertoire of past sounds (not excluding those of speech and literature, which are less trustworthy owing to the vagaries of orthography and phonetic changes in language). The study of contrasting musical styles could help to indicate how, during different periods or different musical cultures, people actually listened differently. For the experience of music shows us that different features or parameters seem to characterize each epoch or school: thus Arab music is noted for rhythm and melody, while that of Western Europe—at least over the last three hundred fifty years—has emphasized harmony and dynamics. To have a good ear, to have musicianship in any culture, means then to have proficiency in selected areas, and the ear training exercises of any musical culture determine what they will be.

A cross-cultural study of the relationship between musical expression and aural perception has never been undertaken, but it should not too long be delayed. It would be of great value in answering questions like these: how does a society regard the relationship between frequency, time and intensity? between continuity and interruption? between impact and steady-state sounds? between foreground and background? signal and noise? or noise and silence—which is to say, dynamism and rest?

Perspective and Dynamics I will give one example of the complementary development of musical expression and aural perception drawn from a single culture, and will try to show how it has developed into a concrete listening attitude. The dimension I want to consider is dynamics, which has an appropriate visual analogue in perspective. Perspective was introduced into European painting during the fifteenth century, and became the predominant style following the works of Masaccio and Uccello. There is only one ideal point from which a perspective painting may be viewed: viz., the point of view. Perspective fastens the viewer to a position directly before the window of the picture frame.

When Giovanni Gabrieli composed his *Sonata Pian' e Forte* (literally, to be sounded soft and loud), he introduced perspective thinking into Western music. Before this date we have no record of dynamic contrast in music, by which we must not infer that it did not exist, but may deduce

that it had not become an articulated desideratum of performance. Gabrieli's *piano* and *forte* were the first steps toward the quantification of sound level, just as the foot and furlong had earlier quantified space. Just as objects are rank-ordered in perspective painting, depending on their distance from the viewer, so musical sounds are rank-ordered by means of their dynamic emphasis in the virtual space of the soundscape. It is an equally deliberate illusion which centuries of training turned into a habit. The classical Western composer places sounds in high definition before the eye of the ear.

Just as perspective focusing is unique to Western art, the organization of music along various dynamic planes is special to Western music. In fact, it is surprising to learn how absent dynamic nuancing is in many musical cultures. Von Békésy reports that, in his experiments on loudness discrimination,

> . . . one of the subjects was a gypsy violinist. In the early part of the experiment his difference limens were enormous, far out of the range of the other subjects. His pitch limens, however, had about the usual values. After much probing it finally developed that he was paying little attention to loudness changes, and the reason for this was that in gypsy music only the pitch is considered an important variable and the loudness is kept relatively uniform. After this situation was understood by the subject, and deliberate training in loudness perception was carried out, his loudness limens fell to normal values.

The same thing was discovered by Catherine and Max Ellis in their work with Australian aborigines. When they were asked to play softer, they simply stopped playing.

By contrast, the exaggerated dynamic plane of Western music allows the composer metaphorically to move sound anywhere from the distant horizon to the immediate foreground. This implication of enormous space and infinite outreach achieved its most striking expressions in the works of Wagner and Debussy. But the important question now is: can we observe any equivalent to these advanced dynamic practices in Western perceptual habits of gathering the soundscape into perspective formations?

If the reader will re-read some of the quotations from the early parts of this book, he will discover an answer to this question. One further example will have to suffice here, though it is a compelling one, for it comes from a person who, in the course of his work, has had to give much thought to the sonic environment—the sound effects technician.

> Faced with a bewildering medley of sounds, the problem is to select those which will illustrate the scene and the accompanying commentary or dialogue to the best advantage. To this end, I recommend what is known as the "three-stage plan." In explanation it is apt to sound rather restrictive, but all it does, in fact, is to impose certain practical

limits on the number of effects to be included, collectively, in any one scene and to decide the degree of prominence that each shall enjoy.

The "three-stage plan" divides the whole sound-scene (called "scenic") into three main parts. These are: The "Immediate," the "Support," and the "Background." The chief thing to bear in mind is that the "Immediate" effect is to be *listened* to while the "Support" and the "Background" effects are merely to be heard. . . .

The "Support" effect refers to sounds taking place in the immediate vicinity which have a direct bearing on the subject in hand, leaving the "Background" effect to its normal job of setting the general scene.

Take, for example, the recording of a commentary at a fun-fair. The "Immediate" effect would be the commentator's voice. Directly behind this would come the "Support" effects of whichever item of fairground amusement he happened to be referring to, backed, to a slightly lesser degree, by the "Background" effect of music and crowd noises.

The three-stage plan of the radio technician corresponds precisely to the classical layout of the orchestral score with soloist, concertino group and tutti accompaniment. And it corresponds to the dynamic listening plane from foreground to horizon which makes focused listening possible. Furthermore, though the point must be made cautiously, the three-stage plan bears a recognizable resemblance to the figure/ground/field division of the (Western) psychologist.

Many other societies never developed the habit of perspective viewing. The study of Eskimo, Chinese and Byzantine art shows how differently space was perceived by these peoples. The Chinese spread objects out over the entire drawing surface, suggesting broad peripheral vision—the opposite of perspective focusing. More curious is the Byzantine convention of reversed perspective, by which objects were frequently enlarged as they receded in space. The Eskimos, as Edmund Carpenter has shown, would often continue a drawing over the edge of the drawing surface onto the back of the material, considering it part of the same surface. Carpenter writes:

> I know of no example of an Aivilik describing space primarily in visual terms. They don't regard space as static and therefore measurable; hence they have no formal units of spatial measurement, just as they have no uniform divisions of time. The carver is indifferent to the demands of the optical eye; he lets each piece fill its own space, create its own world without reference to background or anything external to it. . . . Like sound, each carving creates its own space, its own identity; it imposes its own assumptions.

Carpenter feels the Eskimo's space awareness is acoustic.

> Auditory space has no favoured focus. It's a sphere without fixed boundaries, space made by the thing itself, not space containing the

thing. It is not pictorial space, boxed-in, but dynamic, always in flux, creating its own dimensions moment by moment. It has no fixed boundaries; it is indifferent to background. The eye focuses, pinpoints, abstracts, locating each object in physical space, against a background; the ear, however, favours sound from any direction.

If Carpenter is right, Eskimo culture provides an example of the reverse situation from the European Renaissance; with the Eskimos acoustic space has influenced and even dominated visual space.

Gestures and Textures We have noted several times how focused listening with its implication of distance separating the listener from the sound event is disintegrating before the sound walls of the modern world. The modern lo-fi soundscape possesses no perspective; rather, sounds massage the listener with continual presence. As the population of sounds in the world increases, soloistic gestures are replaced by aggregate textures. Textures and crowds are correlatives. The daily sight of a swiftly moving crowd must have constituted an effect to which the senses had to adapt at first. Only after a new visual technique had been mastered did crowds cease to be confusing and the city-dweller learn to scan them leisurely for a chance display or an interesting figure. Many of Baudelaire's poems reveal this perceptual habit, which was presumably new in his time. "Amid the deafening traffic of the town"—so begins Baudelaire's sonnet "À une passante"—there arises by chance, out of a throng of pedestrians, a woman who snaps the poet's senses to attention by her beauty.

This has happened to us all. We are not looking for anything and we find it. We are not listening to anything but suddenly, out of the commotion, a sound jumps forward to become a figure. It would be inappropriate to say that this type of "unfocused" listening did not exist in the past, but it is possible to say that the circumstances which encourage it are more present in the textures of the post-industrial soundscape.

The present-day increase in statistical exercises and probability theorizing of all kinds is also a reflection of this crowding, nor is it surprising that, precisely at this juncture in history, statistics has entered music as a technique for composition. Iannis Xenakis describes his theory of composition as stochastic. "Stochastics," he explains, "studies and formulates the law of large numbers." More to the point, Xenakis has drawn his inspiration directly from the observation of the contemporary soundscape. He writes:

> But other paths also led to the same stochastic crossroads—first of all, natural events such as the collision of hail or rain with hard surfaces, or the song of cicadas in a summer field. These sonic events are made out of thousands of isolated sounds; this multitude of sounds, seen as a totality, is a new sonic event. This mass event is articulated and forms a plastic mold of time, which itself follows aleatory and sto-

chastic laws. If one then wishes to form a large mass of point-notes, such as string pizzicati, one must know these mathematical laws, which, in any case, are no more than a tight and concise expression of chains of logical reasoning. Everyone has observed the sonic phenomena of a political crowd of dozens or hundreds of thousands of people. The human river shouts a slogan in a uniform rhythm. Then another slogan springs from the head of the demonstration; it spreads towards the tail, replacing the first. A wave of transition thus passes from the head to the tail. The clamor fills the city, and the inhibiting force of voice and rhythm reaches a climax. It is an event of great power and beauty in its ferocity. Then the impact between the demonstrators and the enemy occurs. The perfect rhythm of the last slogan breaks up in a huge cluster of chaotic shouts, which also spreads to the tail. Imagine, in addition, the reports of dozens of machine guns and the whistle of bullets adding their punctuations to this total disorder. The crowd is then rapidly dispersed, and after sonic and visual hell follows a detonating calm, full of despair, dust, and death. The statistical laws of these events, separated from their political or moral context are the same as those of the cicadas or the rain. They are the laws of the passage from complete order to total disorder in a continuous or explosive manner. They are stochastic laws.

There are times when one sound is heard; there are times when many things are heard. *Gesture* is the name we can give to the unique event, the solo, the specific, the noticeable; *texture* is then the generalized aggregate, the mottled effect, the imprecise anarchy of conflicting actions.

A texture may be said to consist of countless inscrutable gestures. They are like the one-celled bacteria which are perceptible only in masses or cluster formations. Thus the sound events in a texture come to be considered statistically as they are in the countless number of sound level surveys being undertaken by so many modern cities in the modern world where noise pollution has broken out of control.

But for the soundscape researcher, the aggregate should never be confused with the singular, for they are not at all the same thing. The soundscape researcher must always remember Zeno's paradox: "If a bushel of corn turned out upon the floor makes a noise, each grain and each part of each grain must make a noise likewise, but, in fact, it is not so."

The aggregate sound of a texture is not merely a simple sum of a lot of individualistic sounds—it is *something different.* Why elaborate combinations of sound events do not become "sums" but "differences" is one of the most intriguing aural illusions.

In the broad-band texture there is also another aural illusion, for in such a sound other sounds may often be heard. I remember when Bruce Davis and I were working on the composition *Okeanos,* which combines the natural polynoise of the sea with electronic sounds and voices reciting maritime poetry. After many hours of working with the tapes of waves,

we often heard in them other parts of the program, submerged as it were, rising at moments to the level of perception, then being carried off again to oblivion by the cascading waters.

Psychologists are aware of this type of aural illusion. In his fascinating little book *Soundmaking*, Peter Ostwald reports on the effects of playing a recording of a baby crying, masked nine decibels by white noise, to a group of patients in a mental hospital. The listeners heard the baby cry variously as

- "a voice shouting, a man's voice trying to be heard, an agitated sound"
- ". . . someone yelling and echoing"
- "a noisy factory with somebody hammering"
- "tremendous machinery, dynamos and . . . people shouting at each other"
- "a high sound, ayee, ayee, like a trumpet."

The polynoise of the sea resembles the white noise of the laboratory. Thus, no two waves are the same, and even the same wave, played over repeatedly on tape, will continue to yield up new secrets to the imagination at each listening. "You never go down to the same water twice," says Heraclitus.

Many other sounds also seem to have these miraculous powers. The wind, for instance, may even surpass the sea in mischievousness, and as witness we recall the contradictory voices of Typhoeus, the wind god of Hesiod's *Theogony*, quoted in Chapter One. In his *Treatise on Painting*, Leonardo da Vinci comments on ". . . the sound of bells, in whose strokes you may find every word which you can imagine." The same thing has been found to exist in words repeated over and over until they hypnotize the mind, at which point they may give rise to new word-sounds. Such is the function of a mantra. Perhaps the reasons why certain sounds produce aural illusions will never be satisfactorily explained. And perhaps it is just as well that they should not be, for an explanation would reduce their rich attraction as sound symbols.